Urdu (in Arabic script)

Romanization Table

Urdu (in Arabic script)

Letters of the Alphabet

Initial	Medial	Final	Alone	Romanization
ا	ﺎ	ﺎ	ا	omit (see Note 1)
ﺑ	ﺒ	ﺐ	ب	b
ﭘ	ﭙ	ﭗ	پ	p
ﺗ	ﺘ	ﺖ	ت	t
ﭨ	ﭩ	ﭧ	ٹ	ṭ
ﺛ	ﺜ	ﺚ	ث	s̱
ﺟ	ﺠ	ﺞ	ج	j
ﭼ	ﭽ	ﭻ	چ	c
ﺣ	ﺤ	ﺢ	ح	ḥ
ﺧ	ﺨ	ﺦ	خ	k͟h
د	ﺪ	ﺪ	د	d
ڈ	ﮉ	ﮉ	ڈ	ḍ
ذ	ﺬ	ﺬ	ذ	ẕ
ر	ﺮ	ﺮ	ر	r
ڑ	ﮍ	ﮍ	ڑ	ṛ
ز	ﺰ	ﺰ	ز	z
ژ	ﮋ	ﮋ	ژ	zh
ﺳ	ﺴ	ﺲ	س	s
ﺷ	ﺸ	ﺶ	ش	sh
ﺻ	ﺼ	ﺺ	ص	ṣ
ﺿ	ﻀ	ﺾ	ض	ẓ
ﻃ	ﻄ	ﻂ	ط	t̤
ﻇ	ﻈ	ﻆ	ظ	z̤
ﻋ	ﻌ	ﻊ	ع	‘ (ayn)
ﻏ	ﻐ	ﻎ	غ	g͟h
ﻓ	ﻔ	ﻒ	ف	f
ﻗ	ﻘ	ﻖ	ق	q
ﻛ	ﻜ	ﻚ	ك	k
ﮔ	ﮕ	ﮓ	گ	g
ﻟ	ﻠ	ﻞ	ل	l
ﻣ	ﻤ	ﻢ	م	m
ﻧ	ﻨ	ﻦ	ن	n
ں	ں	ﮟ	ﮞ	n̲ (see Note 2)
و	ﻮ	ﻮ	و	v

Initial	Medial	Final	Alone	Romanization
ھ	ـھـ	ـہ	ہ	h
-	-	ـة	ة	t (see Rule 10)
یـ	ـیـ (ـے ، ـي)	ـی (ـے ، ـي)	ی	y (see Note 3)

Digraphs Representing Urdu Aspirates (see Note 4)	Value
bh	بھ
ph	پھ
th	تھ
ṭh	ٹھ
jh	جھ
ch	چھ
dh	دھ
ḍh	ڈھ
ṛh	ڑھ
kh	کھ
gh	گھ

Urdu Vowels and Diphthongs (see Note 5)	Value
a	◌َ
u	◌ُ
i	◌ِ
ā	◌َا
á	◌َى ، ىٰ
ū	◌ُو
ī	◌ِی
o	◌و
e	◌ی ، ◌ے
au	◌َوْ
ai	◌َے

Notes

1. For the use of ا (*alif*) to support ء (*hamzah*) and ◌ٓ (*maddah*), see rules 1 and 2, respectively. For the romanization of ء by (*alif*), see rule 12. For other orthographic uses of ا see rules 3-4.

2. For the distinction between ن and ں, see rule 6.

3. For the distinction between ی and ے, see rule 11(c) and (e).
4. For the form of the letter ہ in these digraphs, see rule 9.
5. Vowel points are used sparingly, and for romanization must be supplied from a dictionary.

RULES OF APPLICATION

Letters Which May Be Romanized in Different Ways Depending on Their Context

1. ا (*alif*), و and ی are used to support ء (*hamzah*); see rule 12. When so used, these letters are not represented in romanization.

2. ا (*alif*) is used to support ٓ (*maddah*); see rule 13. When so used, it is not represented in romanization.

3. ا (*alif*) is used after a consonant to indicate the long vowel romanized *ā*.

rāj	راج
karnā	كرنا

In some words of Arabic origin this *alif* appears as a superscript letter over ی representing the *alif maqṣūrah*.

da'vá	دعویٰ

The *alif* is sometimes omitted in writing. It is always represented in romanization.

'Abdurraḥmān	عبد الرحمن ، عبد الرحمان

4. ا (*alif*) may be used as an orthographic sign without phonetic significance. In these cases it is not represented in romanization. See rule 16.

'amlan	عملاً

5. ط appears as a superscript letter over ت, د, and ر when the latter represent the cerebral sounds romanized *ṭ*, *ḍ*, and *ṛ*, respectively.

6. Regardless of pronunciation, undotted forms of the letter ن are romanized *n̲* and dotted forms are romanized *n*.

jahān̲	جہاں

7. و is used:

(a) To represent the consonant sound romanized *v*.

dev	ديو
vujūd	وجود

In some words of Persian origin this consonant, though written, has ceased to be pronounced. It is retained in romanization.

k̲h̲vīsh	خويش

(b) To represent the long vowel romanized *ū*.

Urdū اردو

(c) To represent the long vowel romanized *o*.

os اوس

dost دوست

For the romanization of the conjunction و as *o* see rule 19.

(d) To represent the diphthong romanized *au*.

aur اور

qaumī قومی

(e) To support ء (*hamzah*). See rule 12.

For the use of ّ (*shaddah*) with و see rule 14.

8. ہ is used to represent the consonantal sound romanized *h*.

ham هم

gāh گاہ

Final ہ, though not pronounced, is normally retained in romanization.

kih کہ

guldastah گلدستہ

Exception is made in the case of words whose final syllable ends in an aspirated consonant. When final ہ is added to the letter ھ in this position, it is not represented in romanization.

mukh مکھہ

9. ہ (usually written in the form ھ) is used to represent the aspirated element of the sounds romanized *bh*, *ph*, *th*, *ṭh*, *jh*, *ch*, *ḍh*, *ṛh*, *kh*, *gh*.

phūl پھول

acchā اچھا

For the writing and romanization of words ending in an aspirated consonant, see rule 8.

10. ة and ت, which are sometimes used interchangeably, are both romanized *t*.

ḥikmat حکمة ، حکمت

11. ی is used:

(a) To represent the consonant romanized *y*.

siyāsat سیاست

diyā دیا

(b) To represent the long vowel romanized *ī*.

taṣvīr تصویر

(c) To represent the long vowel romanized *e*.

sher شیر

nevā نیوا

When ی with this value is final, the form ے generally replaces ی.

se سے

laṛke لڑکے

(d) To represent the long vowel romanized *á*. See rule 3.

da‘vá دعوی

‘uqbá عقبی

(e) To represent the diphthong romanized *ai*.

maidān میدان

bail بیل

When ی with this value is final, it is generally written ے .

hai ھے

(f) To support ء (*hamzah*). In this position ی is usually undotted. See rule 12.

For the use of ّ (*shaddah*) with ی see rule 14.

For the use of ی in a *muẓāf* see rule 17.

Romanization of Orthographic Symbols Other Than Letters and Vowel Signs

Although vowel signs are frequently omitted in printed texts, they are always taken into consideration in romanization. The rules for other symbols vary.

12. ء (*hamzah*)

(a) In initial position ء is not represented in romanization.

(b) In medial and final position, when ء represents a consonant, it is romanized ’ (*alif*).

mu’min مؤمن

li’e لئے

bhā’ī بھائی

(c) When ء represents the connective syllable joining a *muẓāf* to what follows, it is romanized *-yi*. See rule 17.

malikah-yi Inglistān ملکۂ انگلستان

13. ٓ (*maddah*)

(a) At the beginning of a word, or following the Arabic article ال, آ is romanized *ā*.

āb آب

(b) At the beginning of a syllable within a word, آ is romanized *ʼā*.

mirʼāt	مرآت
Qurʼān	قرآن

(c) ٓ is otherwise omitted in romanization.

14. ّ (*shaddah* or *tashdīd*) indicates the doubling in pronunciation of the letter over which it is written. It is represented in romanization by doubling the letter or digraph concerned.

caccā	چچّا
khaṭṭā	کھٹّا
makkhī	مکّھی

When ّ occurs over و and ی, these letters are regarded as representing consonants. They are romanized *vv* and *yy*, respectively.

quvvat	قوّت
sayyid	سیّد
Zakariyyā	زکریّا

15. ْ (*sukūn* or *jazm*) indicates the absence of a vowel following the letter over which it is written. It is not represented in romanization.

16. *Tanvīn* (written ٌ, ٍ, ً (اً)) is romanized *un*, *in*, *an*, respectively, when it occurs in a word or expression borrowed from Arabic. Otherwise it is not represented in romanization.

fauran	فوراً
ʻamlan	عملاً

Romanization as Affected by Grammatical Structure

17. *Iẓāfat*.

(a) When the *muẓāf*, the first of two words in the grammatical relationship known as, *iẓāfah* ends in a consonant, *-i* is added to it in the romanization.

tārīk͟h-i Hindūstān	تاریخ هندوستان

(b) When the *muẓāf* ends in a vowel or in silent ہ, *-yi* is added.

daryā-yi shor	دریاے شور
zabān-i Urdū-yi muʻallá	زبان اردوے معلی
malikah-yi Inglistān	ملکۂ انگلستان

For the use of ء (*hamzah*) to indicate the *muẓāf*, see rule 12(c).

18. The Arabic article ال is romanized differently depending on the letters and context with which it is associated.

(a) When it is prefixed to a word beginning with a "moon letter" (خ ، ع ، غ ، ف ، ق ، ك ، م ، و ، ه ، ى ، ا ، ب ، ج ، ح) it is romanized *al*.

al-Qur'ān القران

(b) When it is prefixed to a word beginning with a "sun letter" (ز ، س ، ش ، ص ، ض ، ط ، ظ ، ل ، ن ، ت ، ث ، د ، ذ ، ر) the *l* of the article is replaced in romanization by the same letter or digraph as that which begins the following word.

as-sijill السجل

(c) When it occurs before the second element in a name, the vowel of the article is replaced by the final vowel of the preceding word.

'Abdul'azīz	عبد العزيز
'Abdurrashīd	عبد الرشيد
Abūlfaẓl	ابو الفضل
Z̲ūlqarnain	ذو القرنين
Faẓlullāh	فضل اللّٰه

19. The conjunction و, when used to join two closely associated members of a phrase, is romanized *o*.

mā̄l o asbāb مال و اسباب

Otherwise و is romanized *va*.

20. Rules for the capitalization of English are followed, except that the Arabic article *al* is lowercased in all positions.

21. The macron is used with both capital and lowercase letters.

22. The hyphen is used:

(a) To connect a *muẓāf* with the following vowel or syllable. See rule 17.

(b) To connect the Arabic article *al* with the following word. See rule 18.

23. Foreign words in an Urdu context, including Arabic and Persian words, are romanized according to the rules for Urdu.

جناب هیڈ ماسٹر صاحب گورنمنٹ هائی اسکول

Janāb-i Heḍ Māsṭar ṣāḥib-i Gavarnmanṭ Hā'ī Iskūl

For short vowels not indicated in the script, the Urdu vowels nearest the original pronunciation of the word concerned are supplied in romanization.

24. A quotation in another language using the Arabic script is romanized according to the rules for the language concerned.